AWARD NUMBER: W81XWH-16-1-0739

TITLE: Developing a PTEN-ERG Signature to Improve Molecular Risk Stratification in Prostate Cancer

PRINCIPAL INVESTIGATOR: Luigi Marchionni

CONTRACTING ORGANIZATION: Johns Hopkins University
Baltimore MD 21218

REPORT DATE: October 2017

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2017	Annual	30 Sep 2016 - 29 Sep 2017

4. TITLE AND SUBTITLE

Developing a PTEN-ERG Signature to Improve Molecular Risk Stratification in Prostate Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0739

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Luigi Marchionni and Tamara L. Lotan (partnering PI)

E-Mail: marchion@jhu.edu and tlotan1@jhmi.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Johns Hopkins University
1550 Orleans Street
CRB2, Rm 1M52
Baltimore, MD 21231

8. PERFORMING ORGANIZATION REPORT NUMBER

Johns Hopkins University
Baltimore, MD 21231

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Prostate cancer (PCA) is a clinically and genetically heterogeneous and the development of a molecular classification is critical to distinguish lethal from indolent tumors and minimize overtreatment. Genomic alterations of the *PTEN* and *ERG* genes are among the most common in PCA and there is an interest in exploiting these alterations for routine risk assessment. We found that PTEN loss is most strongly associated with PCA death in patients whose tumors do not carry an *ERG* gene rearrangement, suggesting that ERG absence strengthens PTEN loss association with lethal progression. Despite the widely accessible PTEN/ERG molecular classification, our understanding of their biological interaction along PCA progression remains very limited. Hence, in our study we will perform a comprehensive molecular profiling of well-annotated PCA samples in relation to PTEN and ERG status. Our goals are threefold: 1) to confirm that PTEN/ERG double negative tumors are the most aggressive; 2) to characterize the *expression profiles* associated with PTEN and ERG alterations; and 3) to determine whether such *expression profiles* can be used to improve PCA patient stratification into different risk groups.

15. SUBJECT TERMS

Prostate cancer, PTEN, ERG, ETS, MYC, cell cycle, gene expression, Cap Analysis of Gene Expression (CAGE)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	9	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

1. Introduction

Prostate cancer (PCA) is a clinically and genetically heterogeneous and the development of a molecular classification is critical to distinguish lethal from indolent tumors and minimize overtreatment. Recent technological advances have enabled extraordinary insights into molecular changes occurring in PCA and the *PTEN* and *ERG* genomic alterations have emerged as the most common in PCA. Furthermore, we have found that PTEN loss is associated with PCA death most strongly in patients carrying *ERG* rearrangements, hence there is an interest in exploiting such alterations for routine risk assessment. Furthermore, despite the fact that PTEN and ERG molecular classification is widely accessible, our understanding of their interaction during disease progression is very limited, and a molecular signature of PTEN/ERG loss in PCA is still lacking.

To address these issues, we have formed a collaborative, multi-disciplinary team – led by a urologic pathologist and computational biologist with expertise in PCA molecular pathology and cancer genomics – to perform a comprehensive molecular assessment of well-annotated prostate cancers in relation to PTEN and ERG status using existing and novel data. Our objectives are threefold: 1) to confirm that the tumors with loss of PTEN and lacking *ERG* rearrangement are among the most aggressive; 2) to characterize the *expression profiles* associated with PTEN and ERG alterations; and 3) to determine whether these *expression profiles* can improve the way we stratify prostate cancer patients into different risk groups.

Findings from our proposed research have the potential for both immediate and long-term clinical and translational research applicability. First, by analyzing several large clinical cohorts from multiple institutions, we will be able to confirm the performance of these biomarkers in patient risk stratification. Second, we will also be able to assess if and how PTEN/ERG *molecular signatures* correlate with lethal disease risk in comparison to currently available prognostic assays. Third, we expect to identify novel molecular alterations responsible for the distinct clinical and biological behavior of tumors based on PTEN and ERG status. Lastly, we will also generate a wealth of information about the biologic drivers of prostate cancer behavior, which shall then be utilized by the entire PCA research community.

2. Keywords

Prostate cancer, PTEN, ERG, ETS, MYC, cell cycle, gene expression, RNA sequencing, Cap Analysis of Gene Expression (CAGE)

3. Accomplishments

Below are listed tasks, subtasks, and accomplishments for research site 1 coordinated by the initiating PI (Dr. Marchionni). For site 2 research activities please see progress report of the partnering PI (Dr. Lotan).

Specific Aim 1: Validate association of PTEN and ETS status with risk of lethal prostate cancer	**Timeline (Months)**
Major Task 1: Assessing prostatectomy cohorts on multiple tissue microarrays (TMA) for PTEN, ETS, and cell proliferation rate	1-36
Subtask 3: Analysis of immunostaining and in situ hybridization data from Subtask 2	18-30

Progress on Major Task 1 – Subtask 3: This activity has not yet begun

Specific Aim 2: Leverage multi-dimensional public domain data to discover genomic features and signaling pathways associated with PTEN loss in ERG-positive and ERG-negative PCa.	**Timeline (Months)**
Major Task 1: Exploratory analysis of genomics datasets	1-6
Subtask 1: Examine gene expression distributions and identify outliers and other potential problems:	1-6
Major Task 2: Classify tumors based on PTEN, ETS, and MKI67 status.	6-24
Subtask 1: Use the EM-algorithm to classify tumors as positive or negative based on the expression levels of PTEN, ETS family members, and MKI67	6-12
Subtask 2: Compare expression based classification to IHC and in-situ based status obtained in Specific Aim 1	12-30
Subtask 3: Analysis of PTEN and ETS status in cohorts available from GenomeDX and the public domain	12-24
Major Task 3: Comprehensive meta-analysis of differential gene expression programs modulated by PTEN and ETS status in prostate cancer and characterization of their biological and clinical correlates	12-30
Subtask 1: Use generalized linear model to identify genes differentially expressed and differentially modulated by PTEN and ETS in prostate cancer	12-24
Subtask 2: Identification of relevant biological processes and signaling pathways associated with PTEN/ETS molecular signatures in prostate cancer	18-30
Subtask 3: Development and validation of predictive models based on associated with PTEN/ETS molecular signatures in prostate cancer	24-36

Progress on Major Task 1 – Subtask 1: we have performed exploratory data analysis on all clinically annotated prostate cancer datasets available from the public domain and through the collaboration with GenomeDX. We used statistical summaries and data visualizations techniques (*e.g.,* principal component analysis, hierarchical clustering) to identify outliers and unwanted sources of variation in the data, applying appropriate pre-processing procedures and transformations as required.

Progress on Major Task 2 – Subtask 1: We have used the EM-algorithm to classify tumors as positive or negative based on the expression levels of PTEN, ETS family members, and MKI67. Overall, ERG gene expression proved to be bimodal in all datasets analyzed, with nearly perfect concordance with results from IHC and CNV status. On the contrary, PTEN classification based on EM-classification of gene expression proved more challenging, with some degree of variation between datasets (an example is shown in Figure 1 for the MSKCC cohort).

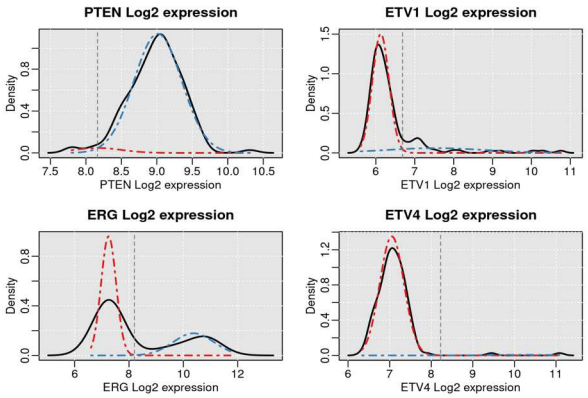

Figure 1: Gene expression distributions for PTEN, ERG, ETV1, and ETV4 in the MSKCC cohort. The underlying distributions from the EM-algorithm are shown in red and blue. ERG and ETV1 expressions are clearly bimodal.

Future plans: In future months, for the patient cohorts for which PTEN/ERG status is known

based on immuno-histochemistry (IHC) and/or copy number variation (CNV) analysis, we will analyze the concordance with the PTEN/ERG status obtained from gene expression using EM-classification. We will further explore alternative methods to classify PTEN status based on gene expression. To this end, we plan to train and validate gene expression based predictors of PTEN IHC and CNV status using the cohorts for which this information is known (TCGA, HPS/HPFS, MSKCC, and JHU cohorts). Finally, using this information, we will proceed with the identification of the molecular signatures associated with PTEN and ERG in prostate cancer.

Progress on Major Task 2 – Subtasks 2 and 3: These activities have just begun.

Training and professional development: Nothing to Report.

Results dissemination to communities of interest: Nothing to Report.

Specific Aim 3: Discover and validate gene regulatory and expression signatures associated with PTEN loss on genetically homogeneous ERG-positive and ERG-negative backgrounds.	**Timeline (Months)**
Major Task 2: Perform CAGE analysis of the tumors resulting from Major Task 1 of Specific Aim3.	6-24
Subtask 1: CAGE library preparation, quality assessment, and sequencing	6-18
Major Task 3: Bioinformatics analysis of CAGE data generated in Major Task 2 of Specific Aim 3.	12-36
Subtask 1: CAGE short reads quality evaluation and alignment to the reference genome	12-24
Subtask 2: Quantification of expressed genomic regions using CAGE tags	18-30
Subtask 3: Classification of expressed genomic regions, identification of active enhancers, promoters, and transcripts	24-30
Subtask 4: Gene expression regulatory network reconstruction and analysis	24-36

Progress on Major Task 2 – Subtask 1: For this task, we have performed initial exploratory analyses to set up the CAGE technology at our institution, performing a pilot experiment as detailed below.

The pilot CAGE experiment was run using four prostate cancer cell lines (two androgen sensitive and two castration resistant) to preserve RNA from tumor samples. The CAGE protocol was performed according to manufacturer's protocol (K.K. DNAFORM) with the following modifications: 1) three micrograms of total RNA were used as input (rather than 5 micrograms); 2) the use of four target samples for simultaneous processing (instead of eight); and 3) the use of Dynabeads for purifications (instead of MPG Streptavidin beads). In additions, adjustments were made for 2nd Strand cDNA synthesis and purification volumes, with final elution after purification at recommended volumes. Resulting yields were low, based on both qPCR and PicoGreen quantitation, but were within expected ranges on a per sample basis if the run would have been performed in the full (8-samples) reaction mode. Quality assessment indicated adequate library preparation, although additional purification was still required.

Future plans: In this experiment, our rationale for the reduced sample processing was an attempt to be conservative with the reagents in the pilot kit (8 reactions), allowing for potential optimization, troubleshooting, and adjustments in a second, follow-up experiment. After

discussing our results with Drs. Carninci and Itoh – the inventors of CAGE at the RIKEN Institute, Yokohama, Japan – however, we realized that by not processing the full recommended group of 8 samples, the optimal efficiency of the protocol was not achieved. We therefore plan to repeat this experiment with the original kit recommendations.

Furthermore, we have also reviewed the nanoCAGE protocol that has been newly developed by our colleagues at RIKEN. This protocol allows total RNA input as low as 50 nanograms. Additionally, the processing time is greatly reduced. Given that the actual experimental samples from tumors may be of lower quantity, we will also evaluate the nanoCAGE protocol with the same full set of cell line samples used for CAGE. The JHSPH Genomics Core laboratory is very experienced and successful in working with RNA and DNA of very low quantity, as well as quality. We are optimistic that we can optimize either protocols to produce the most informative results for this study.

Progress on Major Task 3 – Subtasks 1 through 4: These activities have not yet begun.

Training and professional development: Nothing to Report.

Results dissemination to communities of interest: Nothing to Report.

4. Impact

Impact on prostate cancer research

We have successfully classified ERG status in all available datasets analyzed. Furthermore, we have successfully reproduced in an independent cohort our previous findings indicating that PTEN loss is associated with a worst prognosis in ERG/ETS-negative patients.

We have successfully applied highly validated IHC and in situ hybridization assays to determine PTEN and ETS status in 2 additional cohorts (MSKCC and JHU) with accompanying gene expression data for future analysis.

No molecular signatures of PTEN loss in prostate cancer have been developed to date, thus this project will add significantly to prostate cancer research by further refinement and validation of this prognostic biomarker as we develop expression signatures in the next reporting periods.

Impact on other disciplines: Nothing to Report.

Impact on technology transfer: Nothing to Report.

Impact on society beyond science: Nothing to Report.

5. Changes/Problems

All experiments, and analyses related to Specific Aim 3 Major Task 2 – *i.e.,* setting up the CAGE technology at our institution and perform initial analysis of the tumors collected in Specific Aim 1 – were substantially delayed. Indeed, these activities should have been performed in Dr. Guerrero-Preston laboratory – listed as key personnel in the original proposal – in collaboration with Dr. Marchionni and the Sidney Kimmel Comprehensive Cancer Center (SKCCC) sequencing core. Unfortunately, due to Dr. Guerrero-Preston leaving Johns Hopkins University, however, this task was substantially delayed, and we were forced to consider alternative strategies. As discussed with Dr. Lymor Barnhard, Science Officer for the Prostate Cancer Research Program, we have

searched, within and outside our institution, for other collaborators to perform the analyses and activities related to this Specific Aim. To this end, we have started collaborating with Anne E. Jedlicka, Sr. Research Associate and Manager of the Genomic Analysis and Sequencing Core Laboratory (GASCL) at the Johns Hopkins Bloomberg School of Public Health (JHSPH). A revised budget reflecting these updates is under preparation.

6. Products

Nothing to Report.

7. Participants & Other Collaborating Organizations

Name:	Luigi Marchionni
Project role:	Initiating Principal Investigator
Researcher Identifier:	0000-0002-7336-8071 (ORCID)
Institution:	Johns Hopkins University
Nearest person month worked:	3
Contribution to Project:	Dr. Marchionni coordinated the project, provided supervision of research activities provided by the fellows, and directly performed the analyses
Funding Support:	NA

Name:	Rafael Guerrero-Preston
Project role:	Co-investigator
Researcher Identifier:	NA
Institution:	Johns Hopkins University
Nearest person month worked:	1
Contribution to Project:	Dr. Guerrero-Preston performed some of the initial experiments with CAGE
Funding Support:	NA

Name:	Wikum Dinalankara
Project role:	Post-doctoral fellow
Researcher Identifier:	NA
Institution:	Johns Hopkins University
Nearest person month worked:	2
Johns **Contribution to Project:**	Dr. Dinalankara performed bioinformatics and statistical analyses under Dr. Marchionni supervision
Funding Support:	NA

Name:	Eddie Luidy-Imada
Project role:	Graduate student
Researcher Identifier:	NA
Institution:	Hopkins University
Nearest person month worked:	7

Contribution to Project:	Mr. Luidy-Imada performed bioinformatics and statistical analyses under Dr. Marchionni supervision
Funding Support:	NA

Name:	Ericka M. Ebot
Project role:	Co-investigator
Researcher Identifier:	NA
Institution:	Harvard T.H. Chan School of Public Health
Nearest person month worked:	1 (rounded to 1)
Contribution to Project:	Dr. Ebot provided analytical support for the PHS/HPHS cohorts
Funding Support:	NA

Change in active other support

Dr. Marchionni:

- No longer supported by 1U54RR023561-01A1 (Ford)
- No longer supported by KKESH (Eberhart) – this award is completed
- No longer supported by W81XWH-12-PCRP-TIA – this award is completed
- No longer supported by R01CA163594 (Sidransky) – this award is completed
- PC141474 (Tomlins and Schaeffer) – this award is now active and moved from pending
- R01 PA-13-302 (Marchionni) – this award is now active and moved from pending
- R21 AI124776-01 (Romerio) – this award is now active and moved from pending
- R01CA206027 (Sidransky/Hoque) – this award is now active and moved from pending
- R01CA208709 (Sidransky/Hoque) – this award is now active and moved from pending
- W81XWH-16-PCRP-IDA (Lupold) – this award is now active and moved from pending

Dr. Guerrero-Preston: No longer supported by this award since he left Johns Hopkins University

Other organizations were involved
Organization Name: Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Partner's contribution to the project
Collaboration: Dr. Ericka Ebot provided analytical support for the PHS/HPHS cohorts (<1 person/month effort).

8. Special Reporting Requirements

This project (W81XWH-16-1-0739) is a collaborative award with Dr. Tamara Lotan (Partnering PI, award W81XWH-16-1-0737).

9. Appendices

None